

A	CGGGGGGAGTGGGGAGGAGGGGGTGGCCCGCAGGCATGGAGGCCAATGGACCCGGTCTGAGTCAGAAACCGGTGCTTCAAGGCCACGAAGCATCCCAT M E A N W T A F L F Q A H E A S H 17	90
	CCCCAACAGCAGGCAGCGCAGAACAGCTGCTGCCCTCTGAGTTCTGCTGGAGCCCCCTGATCAGAAACCGGTGCTTCAATACCA H Q Q Q A A Q N S L L P L L S S A V E P P D Q K P L L P I P 47	180
	ATTACTCAGAACCTCAGGTGACAGAACATTAAGGATGCCATTGGGATTAAGAAAAGAAAAACCCAAACTTCGTTGTGCACT I T Q K P Q A A P E T L K D A I G I K K E K P K T S F V C T 77	270
	TACTGCAGTAAAGCATTCAAGGACAGCTATCACCTGAGGCCATCAGCTGCCACACAGGGATCAAGTGGTGTCTGGCAAAGAAA Y C S K A F R D S Y H L R R H O S C H T G I K L V S R A K K 107	360
	ACCCCCACACGGTGGTCCCTTATCTCACCATGGTGGGACAGCAGGCCAATCGTGGTGGTTCAACTATTGAGGCATCTGTCA T P T T V V P L I S T I A G D S S R T S L V S T I A G I L S 137	450
	ACAGTCACTACATCTCCTCGGCCACCAACCCAGCAGCAGCCTAGTACCCAGCAATGCCGTGCCCCAGTCTGCAAGAAACCCAGT T V T T S S S G T N P S S S A S T T A M P V P Q S V K K P S 167	540
	AAGCCTGTCAAGAAGAACACGGCTGTGAGATGTGGAGGGCTTCCGGGATGTGTACCCACTCAATGCCACAAGCTCTCCATTG K P V K K N H A C E M C G K A F R D V Y H L N R H K L S H 197	630
	GACGAAAAGCCCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGGTCTTGTGG D E K P F E C P I C N O R F K R K D R M T Y H V R S H E G G 227	720
	ATACCAAACCCATACTGCACTGTTGTGGAAAGGCTTCTCAAGGCCCTGACCCACTTAAGCAGACATGGGAGGGTCAACAA I T K P Y T C S V C G K F S R P D H L S C H V K H V H S T 257	810
	GAAGACCCCTCAATGCCAACAGTGCACGTGCTGCCATTGGCACCAAGACAGACTACGGACACACATGGTGGCCACGAAGGCAAGGTA E R P F K C O T C T A A F A T K D R L R T H M V R H E G K V 287	900
	TCATGTAACATCTGTGGGAGCTTGTGGTCTGAGCATATATCACCGCCACTTAAAGACACATGGGAGGCCAAAGTATCAACTGTAAAC S C N I C G K L L S A A Y I T S H L K T H G Q S Q S I N C N 317	990
	ACGTGCAAACAAGGCATCAGCAAACGTGCATGAGTGAGGAGGCCAGCAATCAGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG T C K Q G I S K T C M S E E T S N Q K Q Q Q Q Q Q Q Q Q Q Q 347	1080
	CAACAACAACAACATGTGACAAGCTGGCAGGGAGCAGGTAGAGACACTGAGACTGTGGGAGGAAGCTGCAAGCAAGAAAGAA Q Q Q Q Q H V T S W P G K Q V E T L R L W E A V K A R K K 377	1170
	GCTGCCAACCTGTGCCAACCTCCACGGCTGCTACGACACCAGTGACTCTCACTACTCCATTCAATATAACGCTCTGTGTCTGG A A N L C Q T S T A A T T P V T L T T P F N I T S S V S S G 407	1260
	ACTATGTCACACCCAGTCACAGTGGCAGCTGCAATGAGCATGAGAAGTCCAGTAATGTCAGTGCAGTTAACATAACGCCCCCTTA T M S N P V T V A A A M S M R S P V N V S S A V N I T S P L 437	1350
	GCCATGACCTCACCTTAAACACTCACCAACCCAGTCACCTCCCCACCCCTGTGACCGCCCCAGTGAATATAGCACACCCGTCAACC A M T S P L T L T T P V N L P T P V T A P V N I A H P V T I 467	1440
	ACATCTCAATGAACCTGCCACTCTATGACATTAGCTGCCCTCTCAATATAGCAATGAGGCCGTGAGAAAGTATGCCTTGTGCC T S P M N L P T P M T L A A P L N I A M R P V E S M P F L P 497	1530
	CAAGCTTGCCTACGTCAACGCCCTGGTAAACAGTATTATAAGCTAAATGGTTAAAGTAATTTACCAAGCAACTAACCTTAGTT Q A L P T S P P W	1620
	GATTAAGCAAAAGCAGACTATGAAATTGGGGTTTATTATGTTAGTTAAAGAGTGTAGTAGCTCAATTGGTGGGGTTGTT AAAGTAGGGTATATGTAACCTACTGGACCACTTTAGTTACTCAGAAACCCCTTGTGACACCATTGCTAAACAGGATAGTA 1710 1800 1890 1980 2070 2160 2250 2340 2430 2520 2610 2700 2790 2880 2970 3060 3150 3240 3330 3420 3510 3600 3645	1710 1800 1890 1980 2070 2160 2250 2340 2430 2520 2610 2700 2790 2880 2970 3060 3150 3240 3330 3420 3510 3600 3645

FIGURE 1.

31200 (Sheet 2 of 20)

FIGURE 2.

Human DB1 DNA and Protein Sequences:

10	20	30	40	50	60
AGCGGGGGGAGTGGGGAGGAGGGGGGTCGGCCGCCGCAGCCATGGAGGCCAACTGGACCG					
			M	E	A
			N	W	T>
70	80	90	100	110	120
CGTTCCCTGTTCCAGGCCATGAAGCTTCCCATCACCAACAGCAGGCAGCACAGAACAGCT					
A	F	L	F	Q	A
H	E	A	S	H	H
Q	Q	Q	Q	A	A
N	S>				
130	140	150	160	170	180
TGCTGCCCTCCTGAGCTCTGCCGTGGAGCCCCCTGATCAGAAACCATTGCTTCCAATAC					
L	L	P	L	L	S>
S	S	A	V	E	P
D	P	P	D	Q	K
L	L	P	L	L	P
I>					
190	200	210	220	230	240
CAATAACTCAGAACCTCAGGGTGCACCAGAACATTAAAGGATGCCATTGGGATTAAAA					
P	I	T	Q	K	P>
K	P	Q	G	A	P
E	T	E	T	L	K
A	D	A	I	G	I
F	I	G	I	K	K>
250	260	270	280	290	300
AAGAAAAACCCAAAACCTCATTGTGTGCACTTACTGCAGTAAAGCTTCAGGGACAGCT					
K	E	K	P	K	T>
T	S	F	V	C	T
S	C	T	Y	C	S
K	A	F	R	R	D
A	F	R	D	S>	
310	320	330	340	350	360
ATCACCTGAGGCGCCACGAATCCTGCCACACAGGGATCAAGTTGGTGTCCGGCCAAAGA					
Y	H	L	R	R	H>
R	H	E	S	C	H
H	T	G	I	K	L
T	G	I	K	L	V
G	I	K	L	V	S
K	L	V	S	R	P
I>					
370	380	390	400	410	420
AAACCCCCACCACGGTGGTCCCTTATCTCTACCATCGCTGGGACAGCAGCCGAACCTT					
K	T	P	T	T	V>
T	V	V	P	L	I
V	P	L	I	S	T
P	L	I	S	T	I
L	I	S	T	I	A
I	S	T	I	A	G
S	T	I	A	G	D
T	I	A	G	D	S
I	A	G	D	S	S
A	G	D	S	S	R
G	D	S	S	R	T>
430	440	450	460	470	480
CGTTGGTCTCGACCATTGCAGGCATTTGTCAACAGTCACTACATCTCCTCGGGCACCA					
S	L	V	S	T	I>
L	V	S	T	I	A
V	S	T	I	A	G
S	T	I	A	G	I
T	I	A	G	I	L
I	A	G	I	L	S
A	G	I	L	S	T
G	I	L	S	T	V
I	A	G	I	L	T
A	G	I	L	S	V
G	I	L	S	T	K
I	A	G	I	L	K
A	G	I	L	S	P>
490	500	510	520	530	540
ACCCCAGTAGCAGTGCACAGCACAGCTATGCCAGTGAACCCAGTCTGTCAAGAAACCCA					
N	P	S	S	S	A>
P	S	S	S	A	S
S	S	S	A	S	T
S	S	S	T	T	A
S	S	S	T	T	M
S	S	S	T	T	P
S	S	S	T	T	V
S	S	S	T	T	T
S	S	S	S	S	K
S	S	S	S	S	K
S	S	S	S	S	P>
550	560	570	580	590	600
GTAAGCCTGTCAAGAAGAACCATGCTTGTGAGATGTGTGGAAAGGCCTTCCGAGATGTGT					
S	K	P	V	K	K>
K	K	N	H	A	C
N	H	A	C	E	M
H	A	C	E	M	C
A	C	E	M	C	G
C	E	M	C	G	K
E	M	C	G	K	A
M	C	G	K	A	F
G	C	G	K	A	R
C	G	K	A	F	D
G	C	G	K	A	V>
610	620	630	640	650	660
ACCATCTCAATCGACACAAAGCTCTCCCATTCAGATGAGAAACCCTTGAGTGTCTATTT					
Y	H	L	N	R	H>
N	R	H	K	L	S
R	H	K	L	S	H
H	K	L	S	H	S
K	L	S	H	S	D
L	S	H	S	D	E
S	H	S	D	E	K
H	S	D	E	K	P
S	D	E	K	P	F
D	E	K	P	F	E
E	K	P	F	E	C
K	P	F	E	C	P
P	F	E	C	P	I>
670	680	690	700	710	720
GTAATCAGCGCTTCAAGAGGAAGGACCGGATGACTTACCATGTGAGGTCTCATGAAGGAG					
C	N	Q	R	F	K>
Q	R	F	K	R	K
R	F	K	R	K	D
F	K	R	K	D	R
K	R	K	D	R	M
R	K	D	R	M	T
K	D	R	M	T	Y
D	R	M	T	Y	H
M	T	Y	H	V	V
T	Y	H	V	R	R
Y	H	V	R	S	H
H	V	R	S	H	E
V	R	S	H	E	G>
730	740	750	760	770	780
GCATCACCAAACCTATACTTGCAGTGTGGAAAGGCTCTCAAGGCCTGACCACT					
G	I	T	K	P	Y>
I	T	K	P	Y	T
T	K	P	Y	T	C
K	P	Y	T	C	S
P	Y	T	C	S	V
Y	T	C	S	V	C
T	C	S	V	C	G
C	S	V	C	G	K
S	V	C	G	K	G
V	C	G	K	G	F
C	G	K	G	F	S
G	K	G	F	S	R
K	G	F	S	R	P
G	F	S	R	P	D
F	S	R	P	D	H>

31200 (Sheet 2 of 20)

31200 (sheet 3 of 20)

FIGURE 2 (cont)

790 800 810 820 830 840
TAAGCTGTCATGTAAAACATGTCCATTCAACAGAAAGACCCCTCAATGCCAAACGTGCA
L S C H V K H V H S T E R P F K C Q T C>

850 860 870 880 890 900
CTGCTGCCCTTGCACCAAGACAGACTGCGGACACACATGGTGCGCCATGAAGGCAAGG
T A A F A T K D R L R T H M V R H E G K>

910 920 930 940 950 960
TATCATGTAACATCTGGAAAGCTCCTGAGTCAGCATACATCACCAGCCACTTAAAGA
V S C N I C G K L L S A A Y I T S H L K>

970 980 990 1000 1010 1020
CTCATGGGCAGAGCCAAAGTATCAACTGTAATACATGTAAACAAGGCATCAGTAAAACAT
T H G Q S Q S I N C N T C K Q G I S K T>

1030 1040 1050 1060 1070 1080
GCATGAGTGAAGAGACCAGTAACCAAAAGCAGCAGCAGCAGCAGCAGCAACAAC
C M S E E T S N Q K Q Q Q Q Q Q Q Q Q>

1090 1100 1110 1120 1130 1140
AACAAACAAACATGTGACAAGCTGCCAGGGAAAGCAAGTAGAAACACTCAGACTGTGGAAAG
Q Q Q H V T S W P G K Q V E T L R L W E>

1150 1160 1170 1180 1190 1200
AAGCTGTTAAAGCAAGGAAGAAAGAAGCTGCTAACCTGTGCCAACCTCCACGGCTGCTA
E A V K A R K K E A A N L C Q T S T A A>

1210 1220 1230 1240 1250 1260
CGACACCTGTGACTCTCACTACTCCATTCACTATAACATCCTCTGTGTCGTCTGAGACTA
T T P V T L T T P F S I T S S V S S E T>

1270 1280 1290 1300 1310 1320
TGTCAAACCCAGTCACAGTGGCAGCTGCAATGAGCATGAGAAGTCCAGTAAATGTTCAA
M S N P V T V A A A M S M R S P V N V S>

1330 1340 1350 1360 1370 1380
GTGCAGTTAACATAACCAGCCCAATGAACATAGGGCATTCTGTAACTATAACCAAGTCAT
S A V N I T S P M N I G H P V T I T S P>

1390 1400 1410 1420 1430 1440
TATCCATGACCTCTCTTAAACACTCACTACCCCCAGTCACACTCCCCACCCCCGTCACTG
L S M T S P L T L T P V N L P T P V T>

1450 1460 1470 1480 1490 1500
CCCCAGTGAATATAGCACACCCCTGTCACCATCACATCTCCAATGAATCTACCCACACCTA
A P V N I A H P V T I T S P M N L P T P>

1510 1520 1530 1540 1550 1560
TGACATTAGCCGCCCTCTCAATATAGCAATGAGACCTGTAGAGAGCATGCCTTCTGCM
T L A A P L N I A M R P V E S M P F L>

3100 (sheet 4 of 20)

FIGURE 2 (cont)

1570 1580 1590 1600 1610 1620
CCCAAGCTTGCCTACATCACCGCCTGGTAAACAGTATTATAAAATCAAAATATGGGTA
P Q A L P T S P P W * >

1630 1640 1650 1660 1670 1680
AAAGTAAATATTACAGCAACTTAACCTTGTGATTAAAGCAAAAGTAAACCATGA

1690 1700 1710 1720 1730 1740
AATTGGGAGATTTATTACATTAGTTAATAAGAGTGTGGTAGCATTTCTCCAATTGG

1750 1760 1770 1780 1790 1800
CTGGGATTATTCAAAGTAGGGTGTGTATGTAACCTATCACTGGACCACCTAGTTAAC

1810 1820 1830 1840 1850 1860
AGAAATTCTTTAGCTGACAACATTGCTAACAGGAGTAGTAGTTGGCAAGATGAAATG

1870 1880 1890 1900 1910 1920
CCAGAATTAAAACCAATCATAAGTAGAACCCACTTCAAAATAAAAAACAGCATTACTAT

1930 1940 1950 1960 1970 1980
TTCTAATCCCAGGAATCACTTATTGTAACACTAGCAGAACTCTCCCTATACAAG

1990 2000 2010 2020 2030 2040
GTGGATGGCTGATTTAACCTGAAATTAAATCCACAGATTGAGAGCTAGTGTAGAATT

2050 2060 2070 2080 2090 2100
GTCTGTGTTATTGTTTATGAGTAAATACATGCATTGTCATAATAAAATGCATTCAG

2110 2120 2130 2140 2150 2160
AGAATATGCATTTACCTTGGGAATATGTTAATTTCAGGCAGCATTCCCTATGGGAAAG

2170 2180 2190 2200 2210 2220
GTGATACCAGCTCTGATATGCAAAGCATATGATAATTATCATTCTAACCAACGTATA

2230 2240 2250 2260 2270 2280
ATAGGGATTGTGACCTGATATTGGAGATGTAATATTGCTCAGCATATTAATCCCGATG

2290 2300
GAATATAGCATTGTAGTTGACTTTT

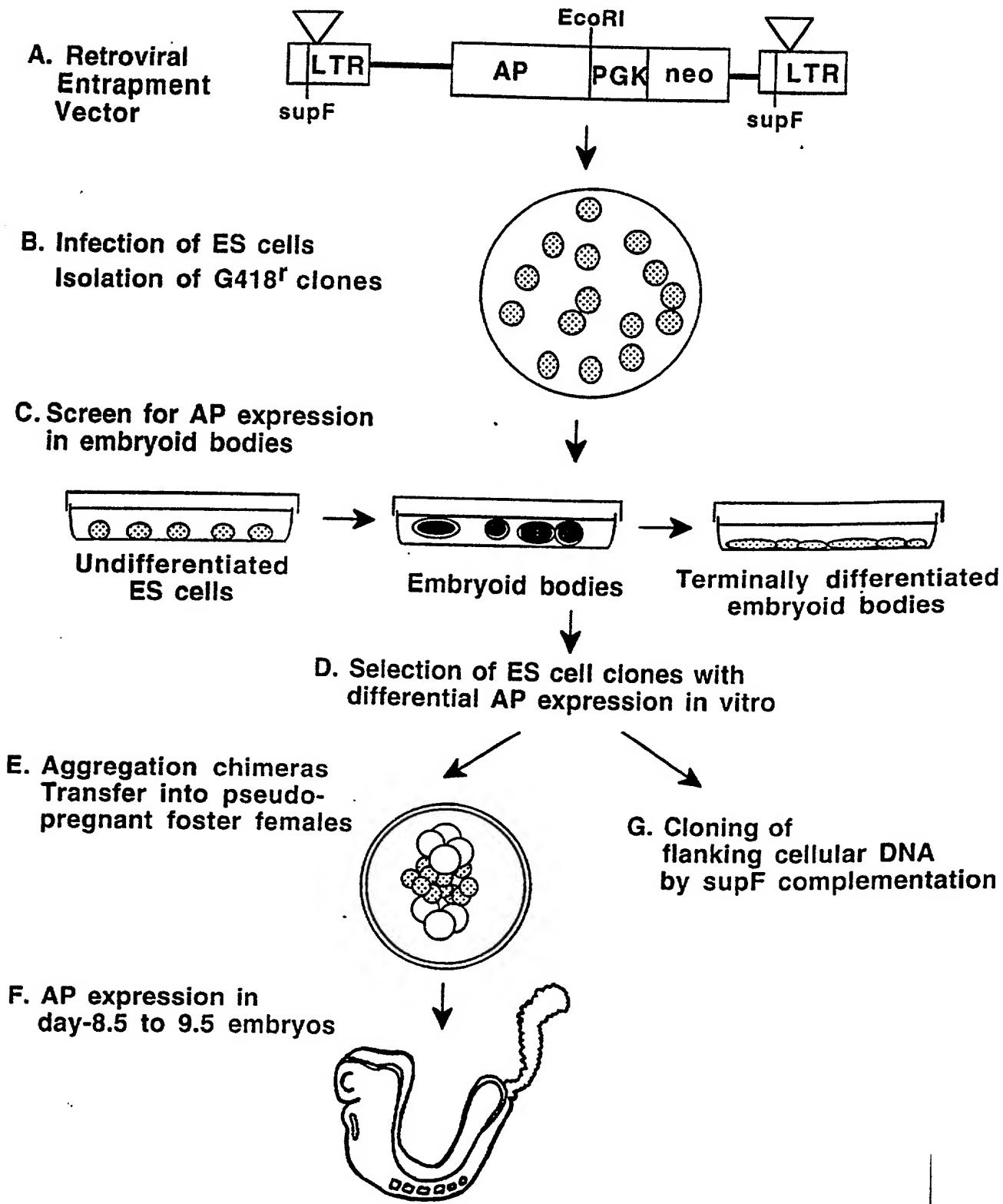


FIGURE 3.

3'200 (sheet 6 of 20)

20270-1625001

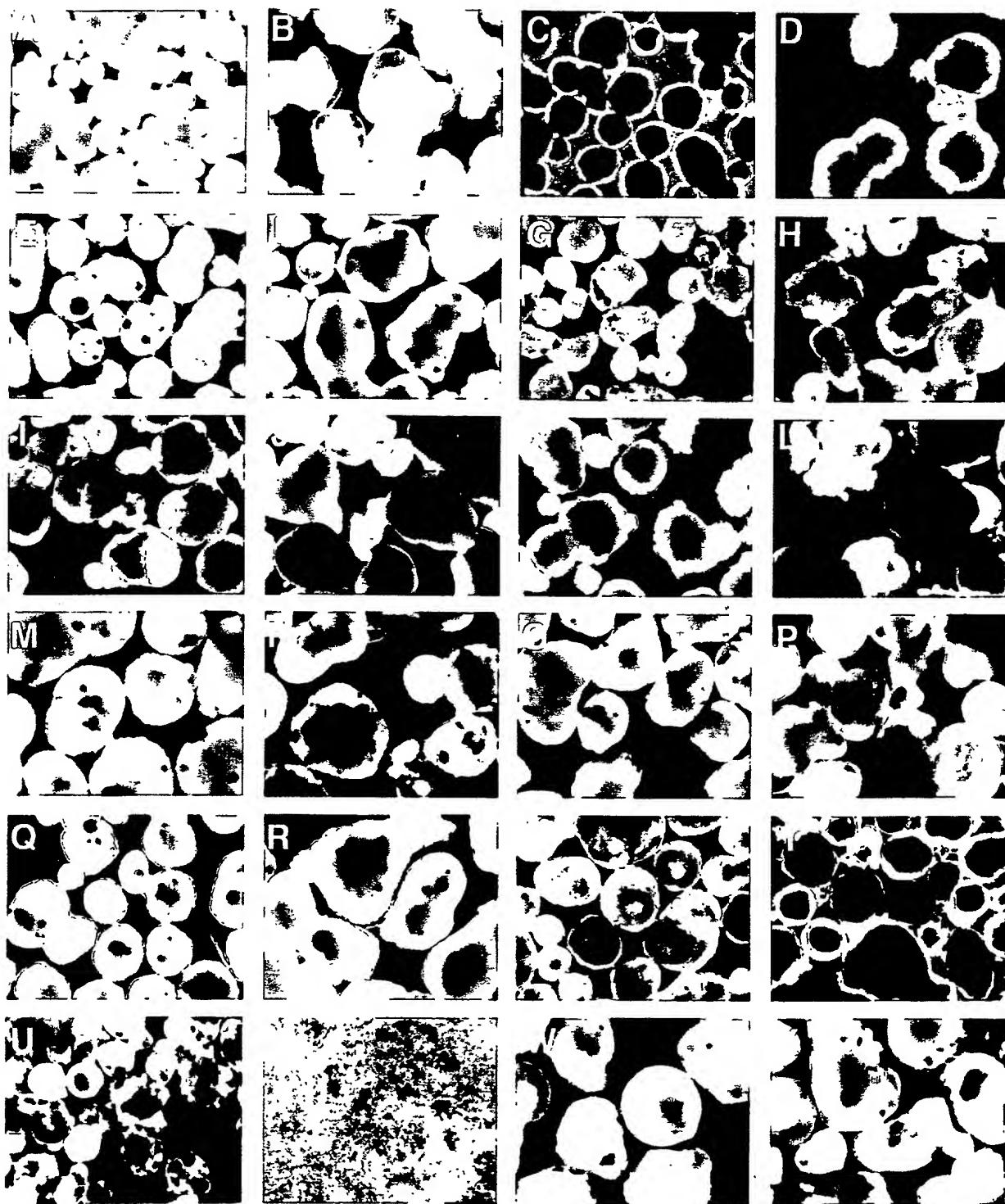


FIGURE 4.

21200 (sheet 7 of 20)

2022700 * 0053667

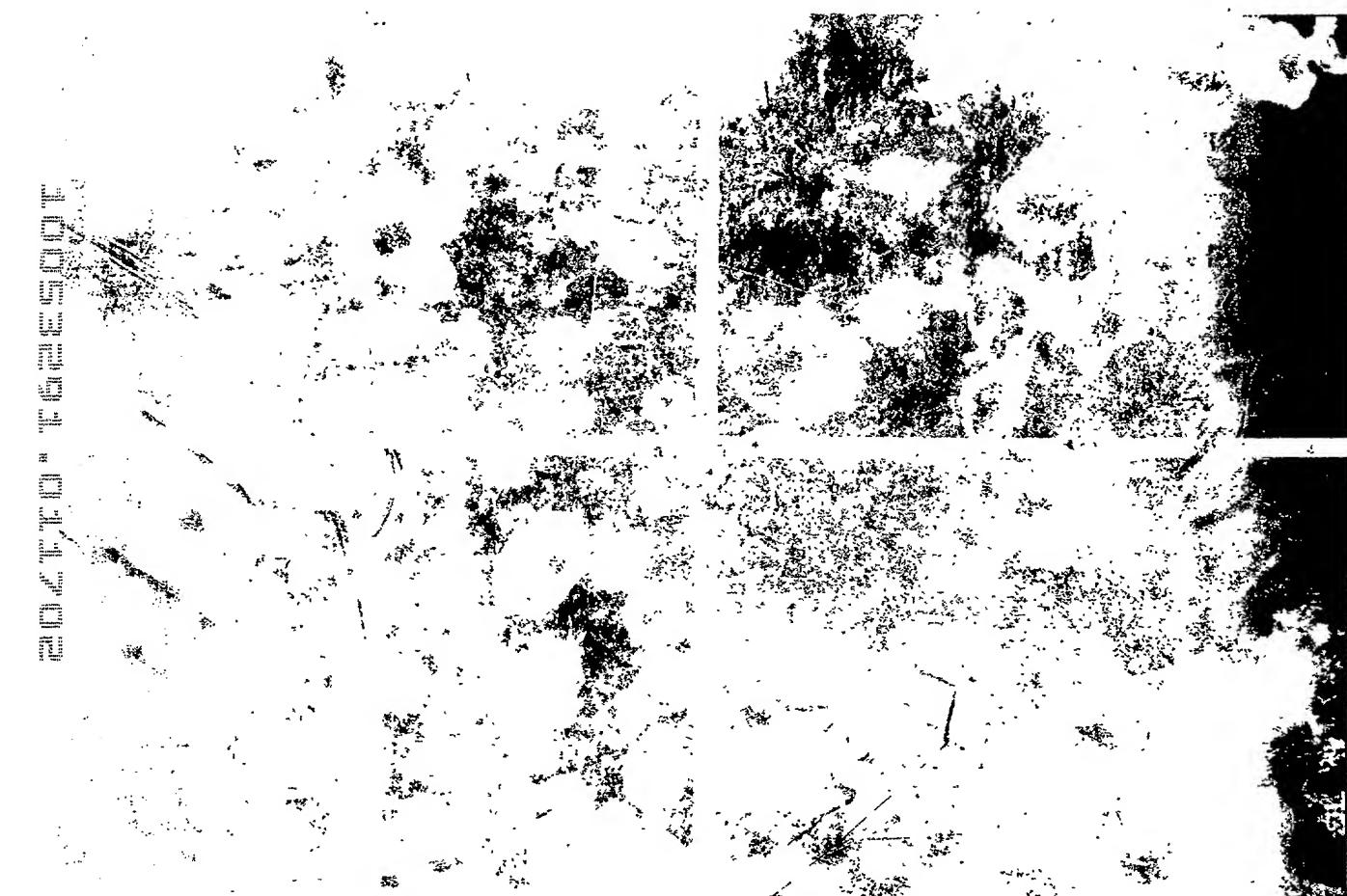
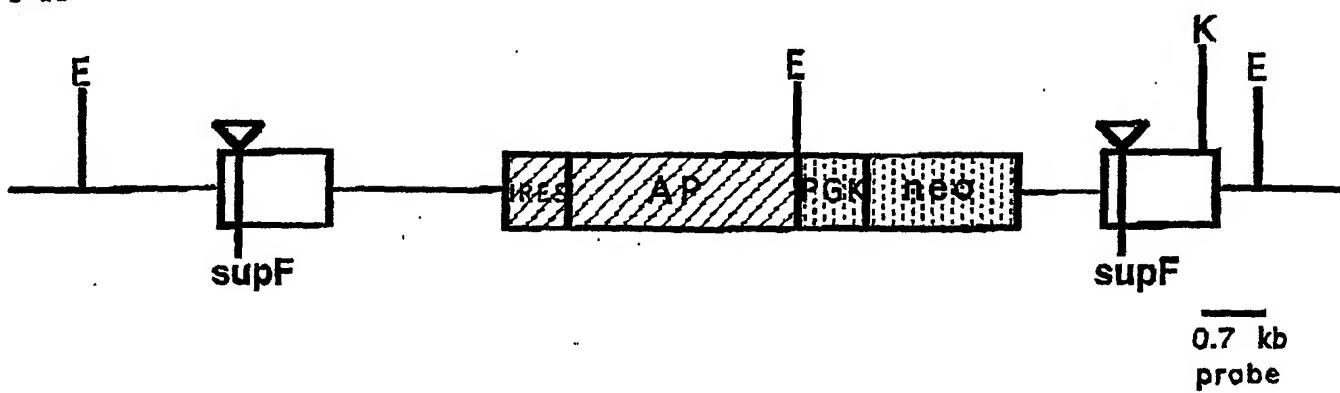


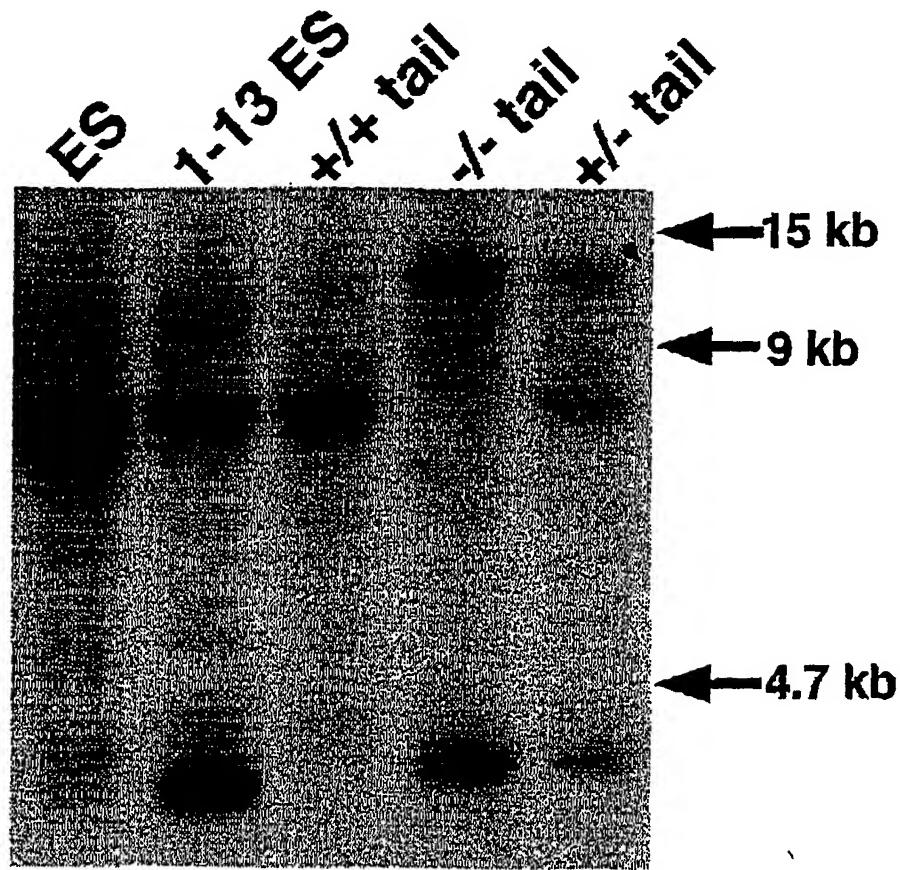
FIGURE 5.

FIGURE 6.

A:



B:



Alignment of *Vezf1/mPur1*:

FIGURE

—

(at 26 fms) 30e/8

1928-1930-1931

3120 (sheet 10 of 10)

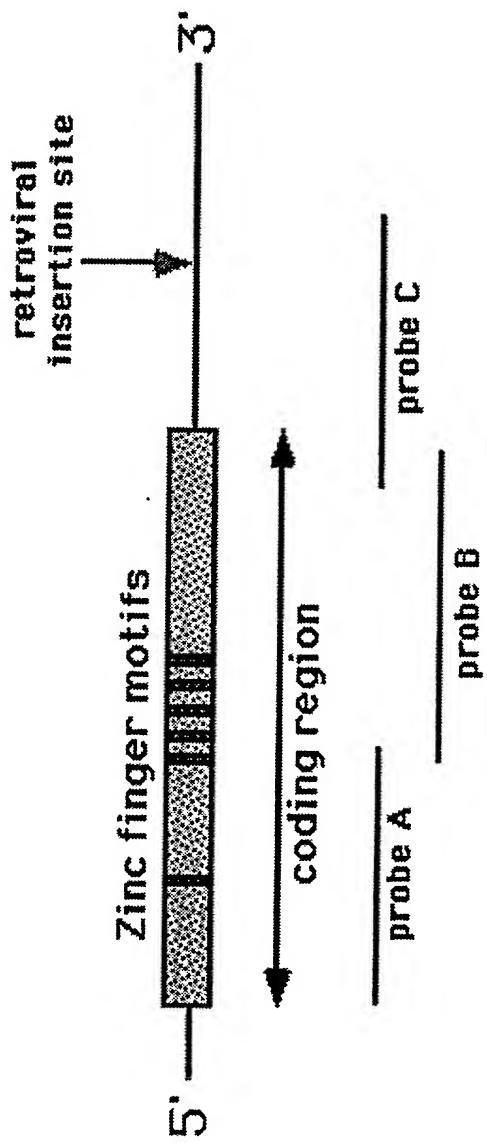


FIGURE 8.

310.00 (sheet 11 of 20)

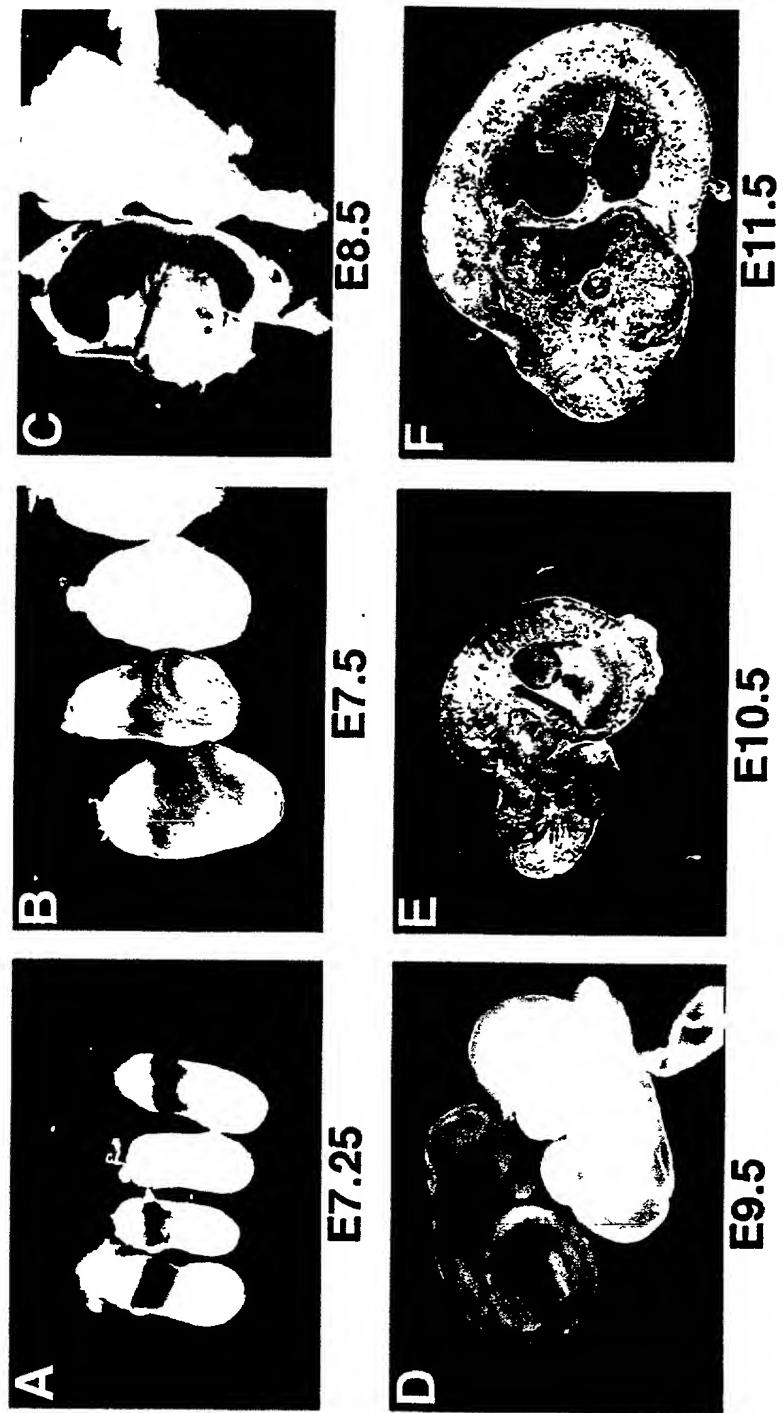


FIGURE 9.

37202 (sheet 2 of 20)

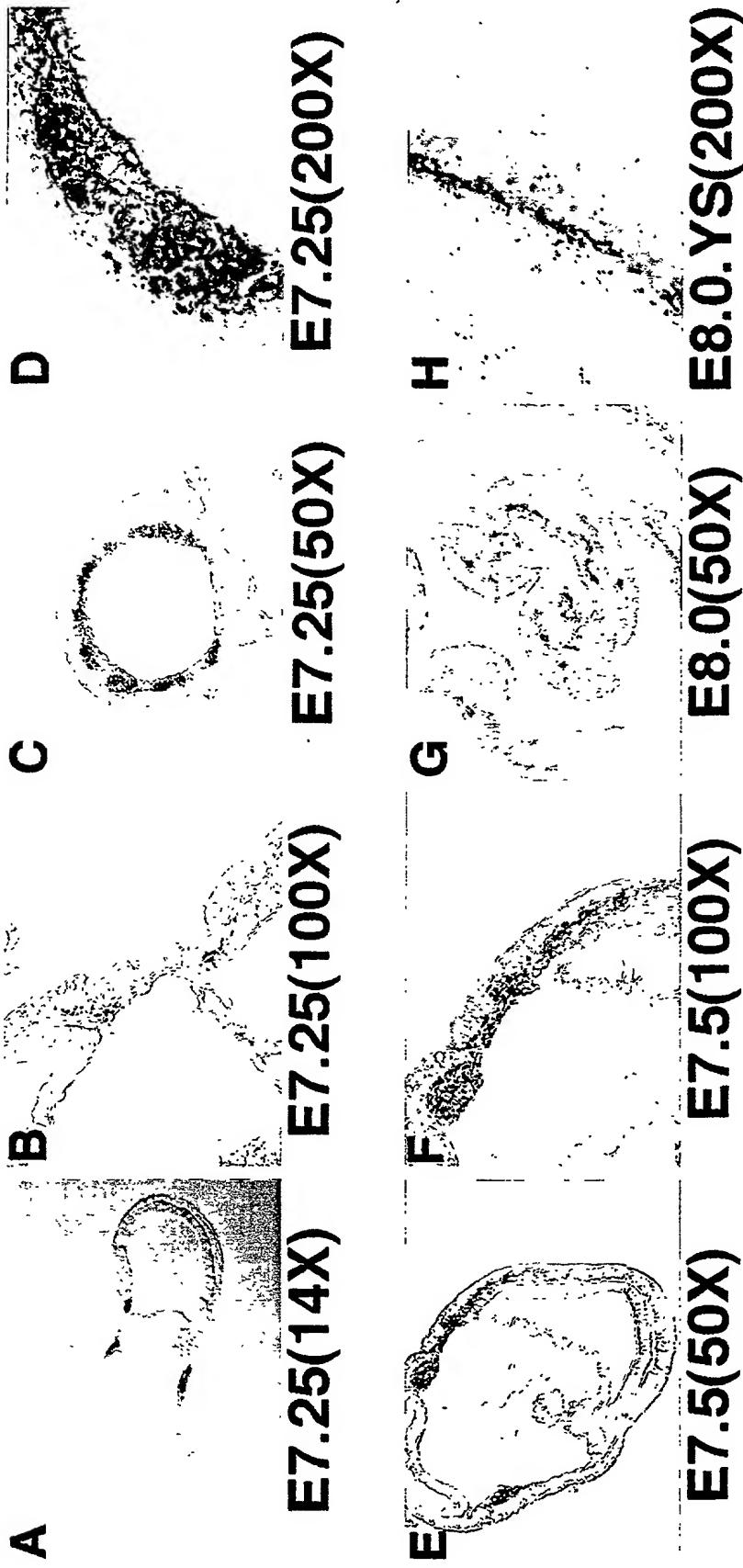


FIGURE 10.

312.0 (sheet 13 of 20)

© 2012 Pearson Education, Inc.

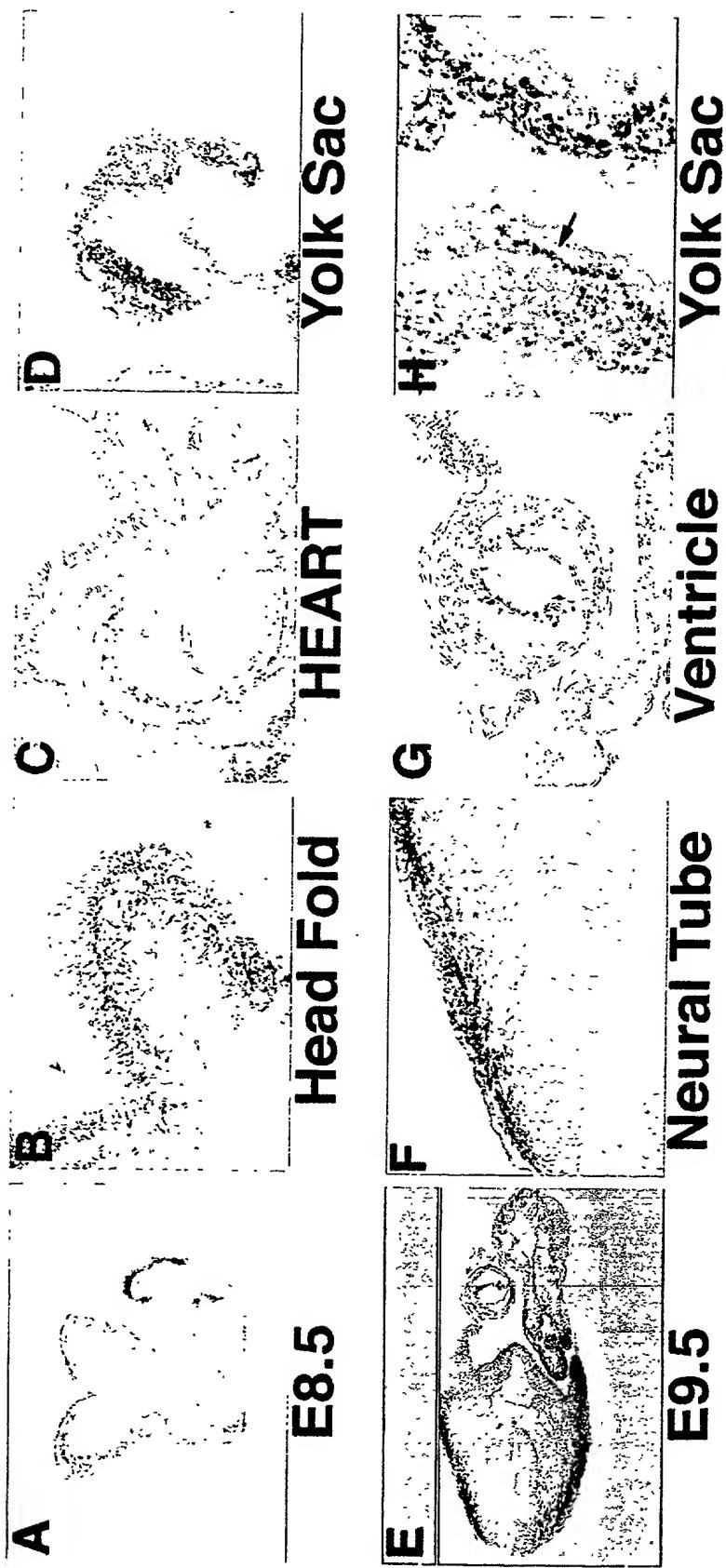


Figure 11.

31.00 (sheet 14 C 2d)

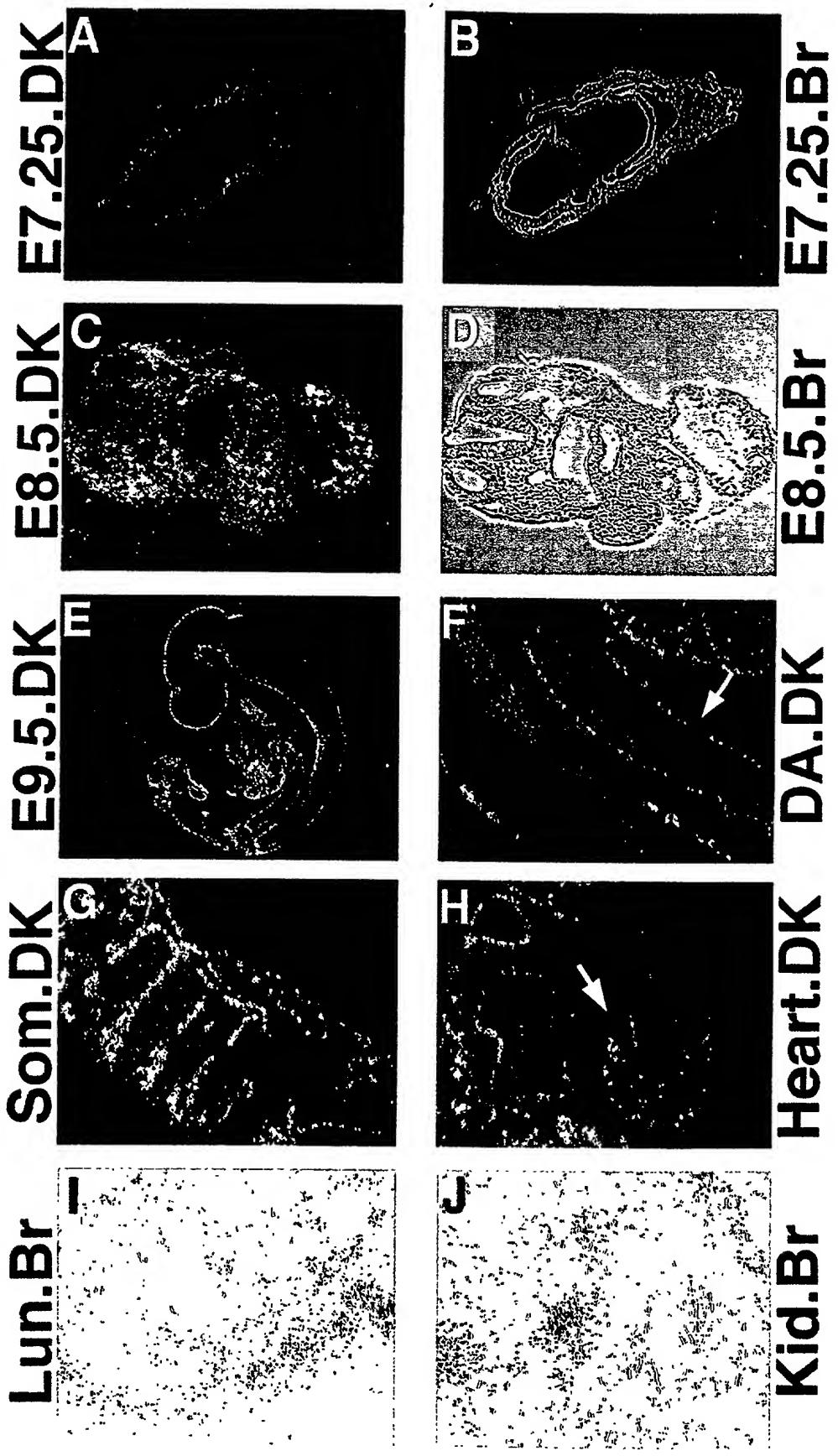


Figure 12.

312.3 (sheet 15 of 20).

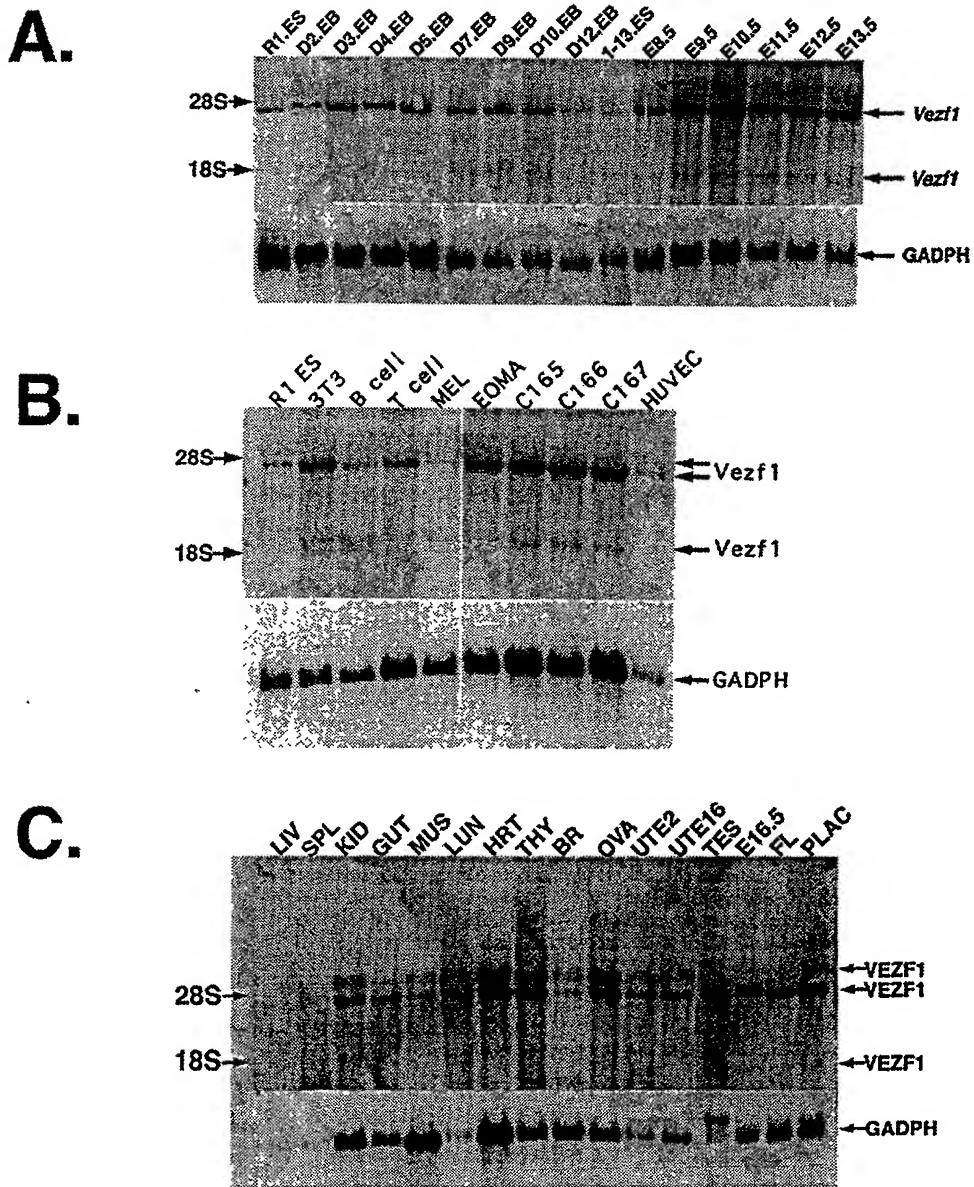


FIGURE 13.

31c30 (Sheet 16 of 20)

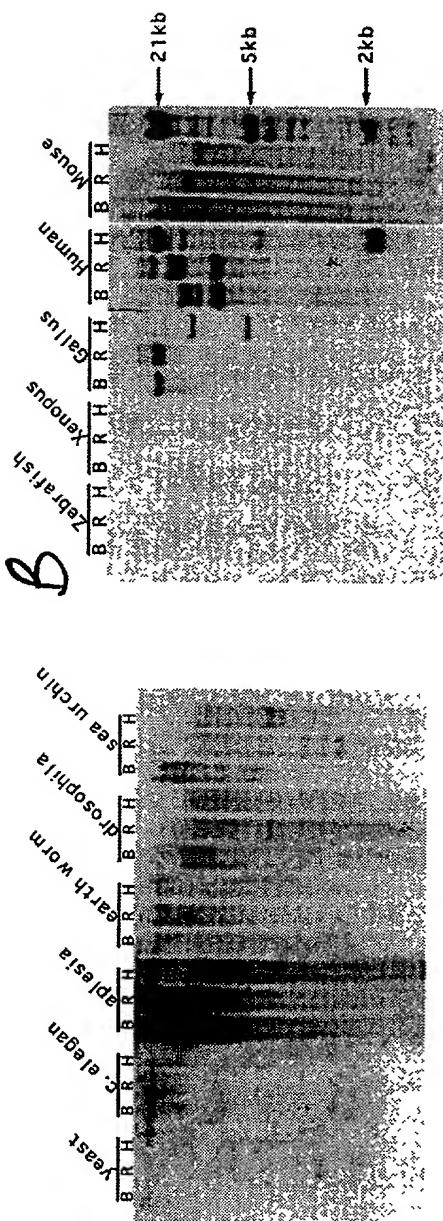
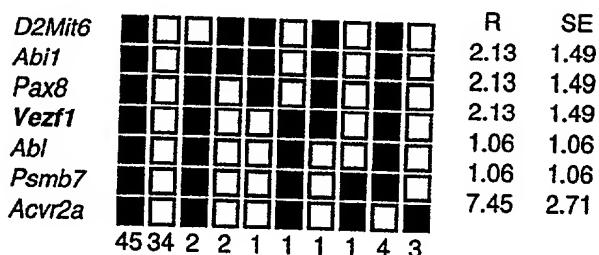


FIGURE 14.

3'200 (sheet 17 of 20)

A: Jackson BSS Chromosome 2



B: Jackson BSS Chromosome 2

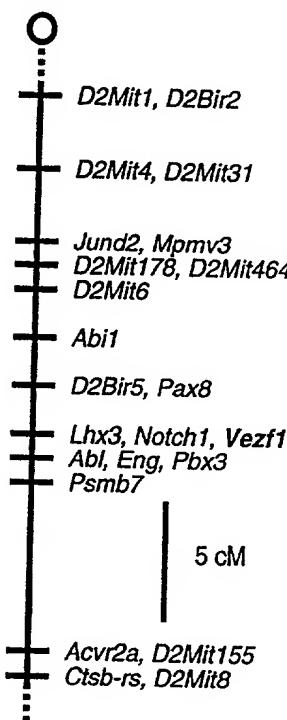
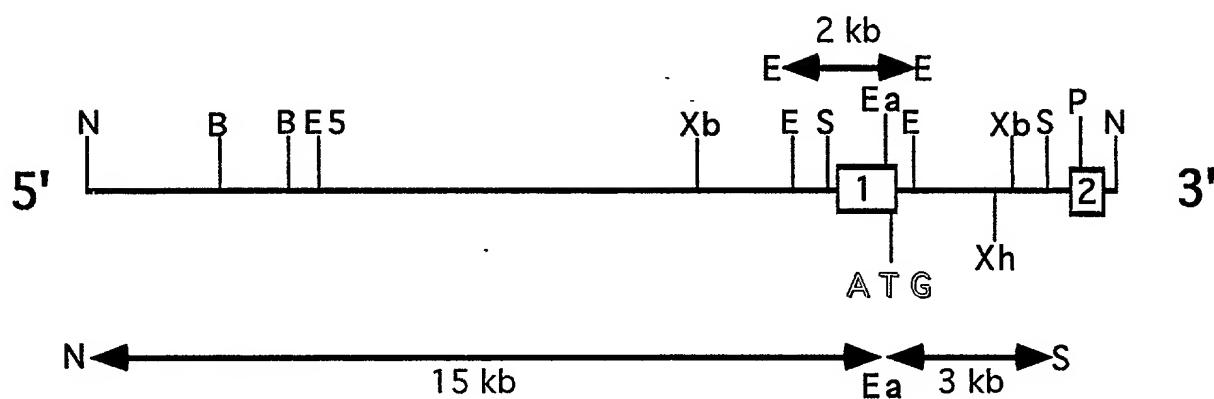


FIGURE
15.

Restriction Enzyme Map of a 20 kb Genomic DNA of the Vezf1 Gene



BamHI (B), EcoRI (E), EcoRV (E5), EagI (Ea), NotI (N), PstI (P), SacI (S), XbaI (Xb), and Xhol (Xh).

— Intronic sequence;

1 Exon 1

2 Exon 2

FIGURE 16.

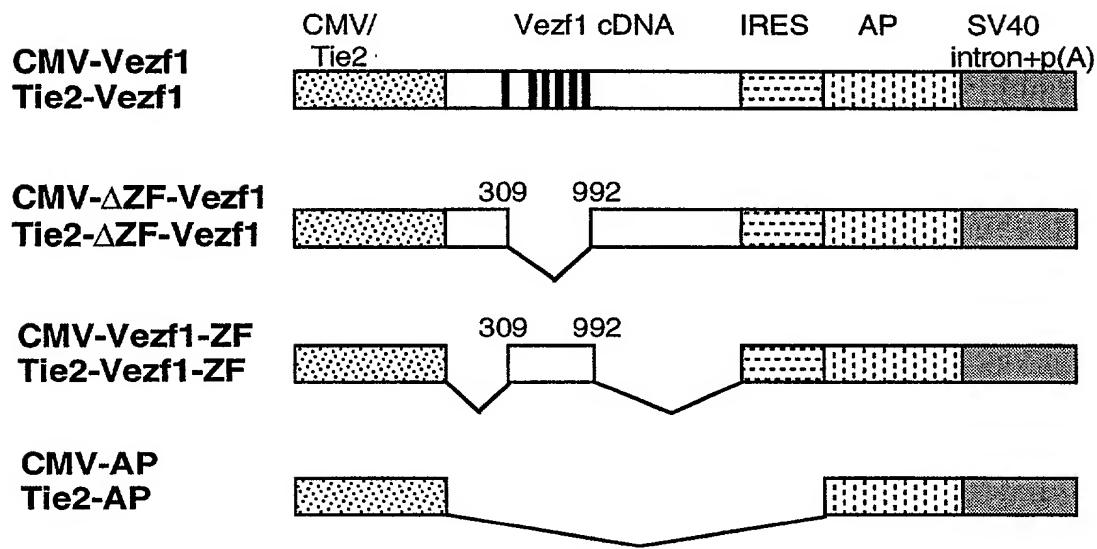
Vezf1 EXPRESSION VECTORS

FIGURE 17.

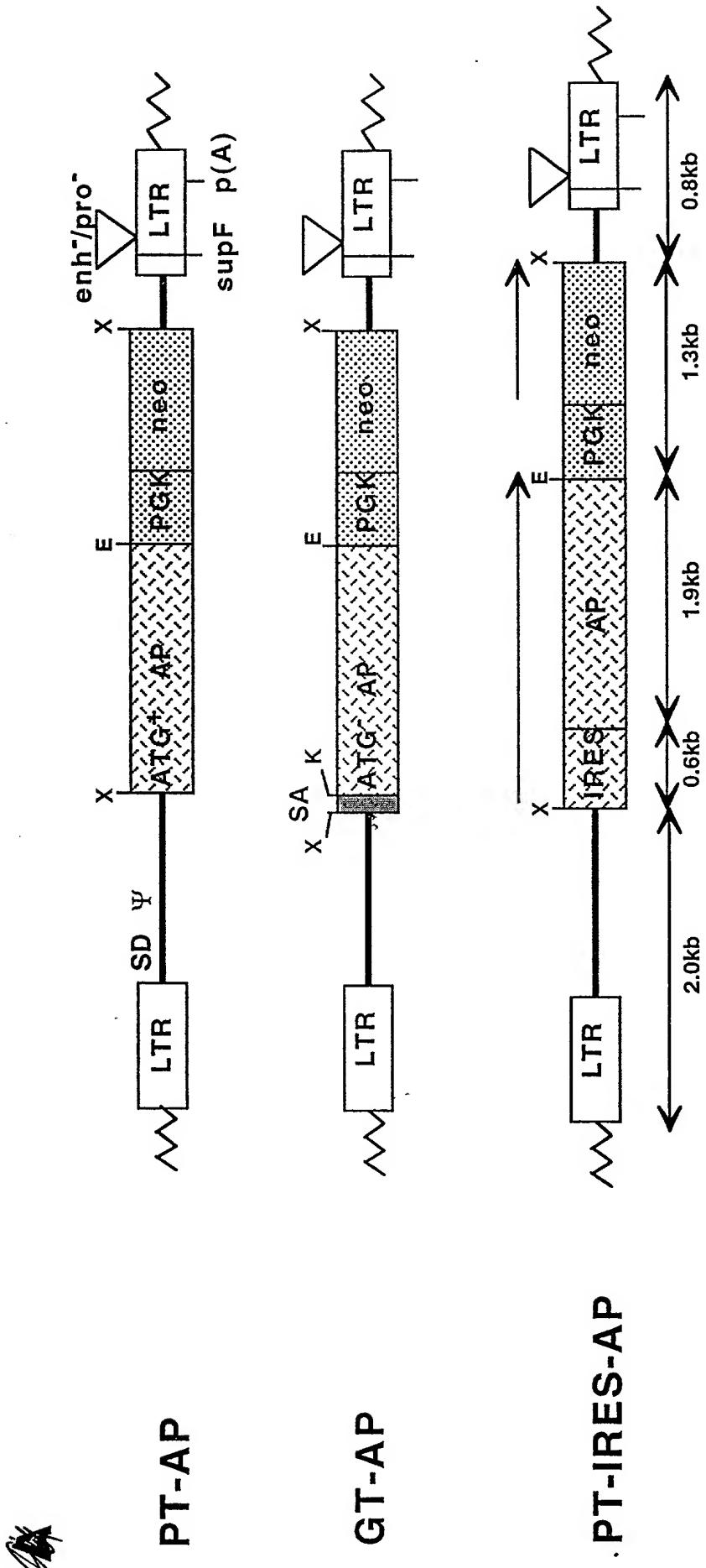


Figure 18.

(22.2 kb + 20.7) 0.715